

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平7-152702

(43)公開日 平成7年(1995)6月16日

(51)Int.Cl. ⁶	識別記号	庁内整理番号	F I	技術表示箇所
G 0 6 F 15/16	3 8 0 Z	8219-5L		
13/00	3 5 5	7368-5B		
H 0 4 L 12/28		7831-5K	H 0 4 L 11/ 00	3 1 0 D

審査請求 未請求 請求項の数4 O L (全 13 頁)

(21)出願番号 特願平5-298644

(22)出願日 平成5年(1993)11月29日

(71)出願人 000003078

株式会社東芝

神奈川県川崎市幸区堀川町72番地

(72)発明者 三浦 勇

東京都青梅市末広町2丁目9番地 株式会

社東芝青梅工場内

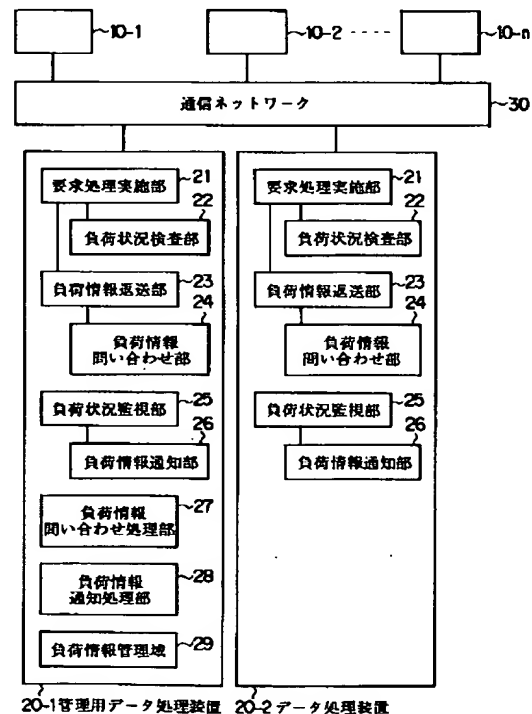
(74)代理人 弁理士 鈴江 武彦

(54)【発明の名称】 分散処理システム

(57)【要約】

【目的】本発明は、端末装置の負荷情報交換上のオーバーヘッド、負荷情報を管理しているデータ処理装置の負荷を削減する。

【構成】負荷情報を生成する負荷状況監視部25と、端末装置からの要求に応じて負荷情報を返送する負荷情報返送部23とを有したデータ処理装置20-1~20-nと、さらに各データ処理装置の負荷情報を処理する負荷情報通知処理部28と、問い合わせに応じて負荷情報を通知する負荷情報問い合わせ処理部27とを有する管理用データ処理装置20-1と、特定の動作時に管理用データ処理装置20-1によって管理された負荷情報に対して問い合わせを行なう負荷情報問い合わせ部13と、負荷情報を記憶するための負荷情報記憶域14と、負荷情報に基づいてデータ処理装置20-1~20-nに対する処理要求の負荷を分散させる負荷分散制御部12とを有する端末装置10-1~10-nを具備して構成する。



【特許請求の範囲】

【請求項 1】 複数の端末装置と、同装置からの処理要求に対応可能な複数のデータ処理装置がネットワークを介して接続された分散処理システムにおいて、

自らの負荷状況を示す負荷情報を生成する負荷状況監視手段と、前記端末装置からの要求に応じて前記負荷状況監視手段によって得られた負荷情報を返送する負荷情報返送手段とを有したデータ処理装置と、

前記複数のデータ処理装置の中の 1 つであって、前記負荷状況監視手段、前記負荷情報返送手段の他に、前記複数のデータ処理装置のそれぞれにおける負荷情報を管理する負荷情報管理手段と、前記負荷情報管理手段によって管理された負荷情報に対する問い合わせに応じて負荷情報を通知する負荷情報問い合わせ処理手段とを有する管理用データ処理装置と、

特定の動作時に前記管理用データ処理装置の前記負荷情報管理手段によって管理された負荷情報に対して問い合わせを行なう第 1 負荷情報問い合わせ手段と、前記第 1 負荷情報問い合わせ手段による問い合わせに応じて、前記負荷情報問い合わせ処理手段によって通知された負荷情報を記憶するための負荷情報記憶手段と、前記負荷情報記憶手段に記憶されたそれぞれのデータ処理装置の負荷情報に基づいてデータ処理装置に対する処理要求の負荷を分散させる負荷分散制御手段とを有する端末装置と、
を具備したことを特徴とする分散処理システム。

【請求項 2】 前記データ処理装置及び前記管理用データ処理装置は、

前記端末装置からの処理要求時に、前記負荷状況監視手段によって検査された負荷状況が所定の状況範囲を越える場合に、前記負荷情報管理手段によって管理された負荷情報を問い合わせる第 2 負荷情報問い合わせ手段をさらに具備し、

前記第 2 負荷情報問い合わせ手段によって得られた負荷情報を前記負荷情報返送手段によって前記端末装置に返送することを特徴とする請求項 1 記載の分散処理システム。

【請求項 3】 前記データ処理装置及び前記管理用データ処理装置は、

前記端末装置からの処理要求に応じた処理の結果と共に、前記負荷情報管理手段によって管理された負荷情報を返送することを特徴とする請求項 2 記載の分散処理システム。

【請求項 4】 前記負荷状況監視手段は、自らの負荷状況を常時監視し、負荷状況が所定の状況範囲に含まれない場合に、負荷状況を示す負荷情報を生成することを特徴とする請求項 1 記載の分散処理システム。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 本発明は、端末装置から複数のデ

ータ処理装置に対して処理要求を行なう分散処理システムに関する。

【0002】

【従来の技術】 従来、複数のデータ処理装置で構成された分散処理システムでは、端末装置がデータ処理装置を選択して処理要求を行なう場合に、端末装置から処理要求を行なう毎に、各データ処理装置の負荷情報を管理している一つのデータ処理装置から負荷情報を得て、その内容に基づき負荷が小さいデータ処理装置を選択していた。

【0003】 この方式を使用した分散処理システムにおける従来の負荷分散制御方式について、図 11 のブロック図で示すシステムを例にして以下に説明する。なお、図 11 に示す分散処理システムは、端末装置 91 から通信ネットワークを介して接続されている複数のデータ処理装置 92 ～ 94 のうちの一つを選択して処理要求を行なうことができる構成となっている。

【0004】 (1) 各データ処理装置 93 ～ 94 は、ある一定の時間間隔、もしくはある一定の負荷状態になったときに、データ処理装置 92 に対し通信ネットワークを介して負荷情報を通知する。

【0005】 (2) データ処理装置 92 は、各データ処理装置 93 ～ 94 からの通知に基づき、各データ処理装置の負荷情報を管理する。データ処理装置 92 自身の負荷情報は、データ処理装置 92 内部の通知手段により同様に管理する。

【0006】 (3) 端末装置 91 は、処理要求を行なう際、先ずデータ処理装置 92 に対して各データ処理装置の負荷情報を要求する。

(4) データ処理装置 92 は、その時点で管理していた各データ処理装置の負荷情報を、端末装置 91 に対して返す。

【0007】 (5) 端末装置 91 は、データ処理装置 92 から得た各データ処理装置の負荷情報に基づき、負荷が小さいデータ処理装置に対して処理要求を行なう。以上のように、分散処理システムにおける従来の負荷分散制御方式では、端末装置 91 が処理要求を行なう毎に、負荷情報を管理しているデータ処理装置 92 に対して、各データ処理装置 92 ～ 94 の負荷情報を問い合わせる情報交換のオーバーヘッドがあった。

【0008】 また、多数の端末装置（図 11 には図示せず）から負荷状態の問い合わせがなされた場合、負荷情報を管理しているデータ処理装置 92 の負荷が増大してしまう。

【0009】

【発明が解決しようとする課題】 このように従来の分散処理システムにおける負荷分散制御方式では、端末装置が処理要求を行なう毎に、負荷情報を管理している 1 つのデータ処理装置に対して、各データ処理装置の負荷情報を問い合わせていたために、負荷情報交換のオーバー

ッドがあった。

【0010】また、多数の端末装置から負荷情報の問い合わせがなされた場合、負荷情報を管理しているデータ処理装置の負荷が増大する等の問題があった。本発明は前記のような事情を考慮してなされたもので、分散処理システムにおいて、端末装置が各データ処理装置の負荷状態を基に負荷分散を行なう際の負荷情報交換上のオーバーヘッド、及び負荷情報を管理しているデータ処理装置の負荷を削減できる分散処理システムを提供することを目的とする。

【0011】

【課題を解決するための手段】本発明は、端末装置が処理要求を行なう際に、従来であれば処理要求の都度、負荷情報を管理しているデータ処理装置に対して、各データ処理装置の負荷情報を問い合わせていたのに対し、負荷情報を管理しているデータ処理装置に対する負荷情報の問い合わせは、端末装置稼働開始時のみとし、その後の処理要求の際は、端末装置自身が記憶している負荷情報を基に負荷分散制御を行なう構成としている。

【0012】そして、本発明においては、端末装置から処理要求を受けたデータ処理装置が、負荷の上限のしきい値を超えている場合、負荷情報を管理しているデータ処理装置から全てのデータ処理装置の負荷情報を得て、端末装置に通知することで、端末装置がデータ処理装置に対して負荷情報を問い合わせなくても、各データ処理装置の負荷状況の変動が、端末装置自身が記憶している負荷情報に反映される構成としている。

【0013】

【作用】このような構成によれば、端末装置が処理要求を行なう際には、端末装置自身が記憶している負荷情報に基づいて負荷分散制御が行われ、各データ処理装置に処理要求が分配される。

【0014】また、端末装置から処理要求を受けたデータ処理装置が、負荷の上限のしきい値を超えている場合、負荷情報を管理しているデータ処理装置から全てのデータ処理装置の負荷情報を得て、端末装置に通知することで、端末装置がデータ処理装置に対して負荷情報を問い合わせなくとも、各データ処理装置の負荷状況の変動が、端末装置自身が記憶している負荷情報に反映される。

【0015】これにより、従来の方式に比べ、端末装置からの負荷情報の問い合わせは、端末装置の稼働開始時以外は発生しない分だけ、負荷情報交換のオーバーヘッドが削減される。また、負荷情報を管理しているデータ処理装置における、負荷情報の問い合わせに係わる負荷も、削減される。

【0016】

【実施例】以下、図面を参照して本発明の一実施例を説明する。図1は、本実施例に係わる分散処理システムの構成を示すブロック図である。図1に示すように、本実

施例の分散処理システムは、複数の端末装置10-1～10-n、複数のデータ処理装置20-1～20-n、通信ネットワーク30によって構成されている。

【0017】端末装置10-1～10-nは、通信ネットワーク30を介してデータ処理装置20-1～20-nの各々と相互に接続されている。また、データ処理装置20-1～20-nは、通信ネットワーク30を介して相互に接続されている。

【0018】端末装置10-1～10-nは、オペレータからの指示に基づきデータ処理装置20-1～20-nに対して処理要求を行なうものである。端末装置10-1～10-nはそれぞれ、データ処理装置20-1～20-nに対して処理要求を発行する処理要求部11と、各データ処理装置に対する処理要求の負荷分散制御を行なう負荷分散制御部12と、管理用データ処理装置20-1に対して各データ処理装置の負荷情報を問い合わせる負荷情報問い合わせ部13と、各データ処理装置の負荷情報を記憶する負荷情報記憶域14とを有している。

【0019】データ処理装置20-1～20-nは、各端末装置10-1～10-nからの処理要求を受けて対応する処理を実施するものである。特にデータ処理装置20-1は、各データ処理装置20-1～20-nに対する負荷情報を管理する構成となっている。以下において、データ処理装置20-1は、他のデータ処理装置20-2～20-nの構成に、さらに各データ処理装置20-1～20-nの負荷情報を管理する機能が設けられた管理用データ処理装置20-1として説明する。

【0020】図2には管理用データ処理装置20-1、データ処理装置20-2の詳細な構成を示している。データ処理装置20-2は、図示せぬデータ処理装置20-3～20-nと同一の構成を有している。

【0021】図2に示すように、データ処理装置20-1～20-nはそれぞれ、端末装置10-1～10-nからの処理要求を受け付けて対応する処理を実施する要求処理実施部21と、要求処理実施部21からの依頼を受けて自データ処理装置の負荷状況を検査する負荷状況検査部22と、各端末装置10-1～10-nからの処理要求に対する応答として各データ処理装置の負荷情報を返す負荷情報返送部23と、管理用データ処理装置20-1に対して各データ処理装置の負荷情報を問い合わせる負荷情報問い合わせ部24と、自データ処理装置の負荷状況を随時監視する負荷状況監視部25と、自データ処理装置の負荷情報を管理用データ処理装置20-1に通知する負荷情報通知部26とを有している。

【0022】さらに、管理用データ処理装置20-1は、前述したデータ処理装置20-1～20-nが有している各部に加え、端末装置10-1～10-n、及びデータ処理装置20-2～20-nからの各データ処理装置の負荷情報問い合わせ要求に対して各データ処理装

置の負荷情報を返す負荷情報問い合わせ処理部 27 と、データ処理装置 20-2~20-n からの負荷情報の通知を処理する負荷情報通知処理部 28 と、各データ処理装置の負荷情報を管理する負荷情報管理域 29 とを有している。

【0023】図 3 はデータ処理装置 20-1~20-n の稼働開始時の動作を説明するためのフローチャート、図 4 は管理用データ処理装置 20-1 がデータ処理装置 20-1~20-n から負荷情報の通知を受けた場合の動作を説明するためのフローチャート、図 5 は端末装置 10-1~10-n の稼働開始時の動作を説明するためのフローチャート、図 6 は管理用データ処理装置 20-1 が端末装置 10-1~10-n もしくはデータ処理装置 20-2~20-n から負荷情報の問い合わせを受けた場合の動作について説明するフローチャート、図 7 は端末装置 10-1~10-n のデータ処理装置 20-1~20-n に対して処理要求を発行する場合の動作について説明するためのフローチャート、図 8 及び図 9 はデータ処理装置 20-1~20-n が端末装置 20-1~20-n から処理要求を受けた場合の動作を説明するためのフローチャート、図 10 はデータ処理装置 20-1~20-n が稼働時に随時実施する負荷状況監視の動作を説明するためのフローチャートである。

【0024】次に、本実施例の分散処理システムにおける負荷分散制御方式の動作について、図 3~図 10 に示すフローチャートを参照しながら説明する。まず、データ処理装置 20-1~20-n が稼働を開始するとそれぞれ、負荷状況監視部 25 によって自データ処理装置の負荷状況を監視し自データ処理装置の負荷情報を生成する (図 3 ステップ A1)。

【0025】データ処理装置が管理用データ処理装置 20-1 の場合 (図 3 ステップ A2)、負荷情報通知部 26 によって自データ処理装置の負荷情報を負荷情報管理域 29 に設定する (図 3 ステップ S3)。

【0026】管理用データ処理装置以外のデータ処理装置 20-2~20-n の場合 (図 3 ステップ A2)、負荷情報通知部 26 によって自データ処理装置の負荷情報を管理用データ処理装置 20-1 に通知する (図 3 ステップ A4)。

【0027】管理用データ処理装置 20-1 は、データ処理装置 20-2~20-n から負荷情報の通知を受けると、通知された負荷情報を負荷情報通知処理部 28 において負荷情報管理域 29 に各データ処理装置毎に設定する (図 4 ステップ B1)。

【0028】次に、端末装置 10-1~10-n が稼働を開始するとそれぞれ、負荷情報問い合わせ部 13 によって管理用データ処理装置 20-1 に対して各データ処理装置 20-1~20-n の負荷情報を問い合わせる (図 5 ステップ C1)。

【0029】管理用データ処理装置 10-1 は、端末装

置 10-1~10-n の何れからより負荷情報の問い合わせを受けると、負荷情報問い合わせ処理部 27 によって各データ処理装置 20-1~20-n の負荷情報を負荷情報管理域 29 から読みだし (図 6 ステップ D1)、問い合わせ元の端末装置に返送する (図 6 ステップ D2)。

【0030】端末装置 10-1~10-n は、各データ処理の負荷情報が管理用データ処理装置 10-1 から返送されると、それらの負荷情報を負荷情報記憶域 14 に記憶する (図 5 ステップ C2)。

【0031】さて、端末装置 10-1~10-n は、オペレータより処理依頼を受けると、先ず処理要求部 11 によってその依頼を受け付ける (図 7 ステップ E1)。端末装置 10-1~10-n は、負荷分散制御部 12 によって負荷情報記憶域 14 を参照し、負荷が小さいデータ処理装置を選択する。もしある一定時間内に複数の処理依頼がなされた場合は、各データ処理装置 20-1~20-n の負荷状況に応じて処理要求数が平均的になるように、処理要求を行なうべきデータ処理装置を選択する (図 7 ステップ E2)。

【0032】そして、端末装置 10-1~10-n は、処理要求部 11 によって、選択されたデータ処理装置に対して処理要求を発行する (図 7 ステップ E3)。データ処理装置 20-1~20-n は、端末装置 10-1~10-n の何れかからの処理要求を受けると (図 8 ステップ F1)、先ず要求処理実施部 21 においてその要求を受け付ける。次に負荷状況検査部 22 において自データ処理装置の負荷状況を調べる (図 8 ステップ F2)。

【0033】負荷が上限のしきい値を超えていない場合は (図 8 ステップ F3)、要求処理実施部 21 において端末装置 10-1~10-n から要求された処理を実施して、その結果を要求元の端末装置に返送する (図 8 ステップ F4)。

【0034】もし、負荷が上限のしきい値を超えている場合は (図 8 ステップ F3)、各データ処理装置の負荷情報を取得する。ここで処理要求を受けたデータ処理装置が管理用データ処理装置 20-1 である場合は (図 9 ステップ F5)、負荷情報問い合わせ部 24 によって、負荷情報管理域 29 から各データ処理装置 20-1~20-n の負荷情報を読み込む (図 9 ステップ F6)。

【0035】一方、その他のデータ処理装置 20-2~20-n である場合は (図 9 ステップ F5)、負荷情報問い合わせ部 24 によって管理用データ処理装置 20-1 に対して各データ処理装置 20-1~20-n の負荷情報を問い合わせる (図 9 ステップ F9)。

【0036】データ処理装置 20-2~20-n から負荷情報の問い合わせを要求された管理用データ処理装置 20-1 は、負荷情報問い合わせ処理部 27 によって各データ処理装置の負荷情報を負荷情報管理域 29 から読みだし (図 6 ステップ D1)、問い合わせ元のデータ処

10

20

30

40

50

理装置に返送する(図DステップD2)。

【0037】このようにして得られた各データ処理装置の負荷情報を、負荷情報返送部23によって処理要求元の端末装置に対する返送情報に設定する(図9ステップF8)。

【0038】そして、要求処理実施例部21において端末装置10-1~10-nから要求された処理を実施して、その結果を先に返送情報に設定した各データ処理装置の負荷情報と共に要求元の端末装置に返送する(図9ステップF9)。

【0039】端末装置10-1~10-nは、データ処理装置20-1~20-nから処理要求に対する応答を得ると、処理要求部11において先ず返送情報に各データ処理装置の負荷情報が設定されているかどうかを調べる(図7ステップE4)。ここで、負荷情報が設定されていない場合は、処理要求に対する処理結果をオペレータに提示する(図7ステップE5)。

【0040】もし、負荷情報が設定されている時は、その内容を負荷情報記憶域14に設定する(図7ステップE5)。そして処理要求に対する処理結果をオペレータに提示する(図7ステップE5)。

【0041】ところで、データ処理装置20-1~20-nは、前述したように、端末装置10-1~10-nから要求された処理を実施するのは独立に、負荷状況監視部25によって随時、自データ処理装置の負荷状況を監視している(図10ステップG1)。

【0042】この監視によって、負荷が上限のしきい値を超えている場合(図10ステップG2)、もしくは下限のしきい値を下回っている場合(図10ステップG3)を検出した際には、負荷状況監視部25によって自データ処理装置の負荷情報を生成する(図10ステップG4)。

【0043】ここで管理用データ処理装置20-1である場合は(図10ステップG5)、負荷情報通知部26において自データ処理装置の負荷情報を負荷情報管理域29に設定する(図10ステップG6)。

【0044】その他のデータ処理装置20-2~20-nである場合は(図10ステップG5)、負荷情報通知部26において自データ処理装置の負荷情報を管理用データ処理装置20-1に通知する(図10ステップG7)。

【0045】データ処理装置20-2~20-nから負荷情報の通知を受けた管理用データ処理装置20-1は、負荷情報通知処理部28によって、負荷情報管理域29の通知元データ処理装置の負荷情報を、通知された負荷情報の内容に更新する(図4ステップB1)。

【0046】このようにして、分散処理システムにおいて、従来であれば端末装置10-1~10-nが処理の要求都度、負荷情報を管理しているデータ処理装置20-1に対して各データ処理装置20-1~20-nを問

い合わせていたのに対し、本発明では負荷情報を管理している管理用データ処理装置20-1に対する負荷情報の問い合わせは、端末装置稼働開始時のみとしている。

【0047】その後の処理要求の際は、端末装置10-1~10-n自身が記憶している負荷情報を基に負荷分散制御を行なう。そして、端末装置10-1~10-nから処理要求を受けたデータ処理装置が負荷の上限のしきい値を超えている場合、負荷情報を管理している管理用データ処理装置20-1から全てのデータ処理装置20-1~20-nの負荷情報を得て端末装置10-1~10-nに通知することで、端末装置がデータ処理装置に対して負荷情報を問い合わせなくとも各データ処理装置の負荷状況の変動が端末装置自身が記憶している負荷情報に反映させることができる。

【0048】従って、負荷情報交換のオーバーヘッドを削減することができる。また、負荷情報を管理している管理用データ処理装置20-1における負荷情報の問い合わせに係わる負荷も削減することができる。

【0049】

【発明の効果】以上詳述したように本発明によれば、端末装置が処理要求を行なう際には、端末装置自身が記憶している負荷情報に基づいて負荷分散制御が行われ、各データ処理装置に処理要求が分配される。また、端末装置から処理要求を受けたデータ処理装置が、負荷の上限のしきい値を超えている場合、負荷情報を管理しているデータ処理装置から全てのデータ処理装置の負荷情報を得て、端末装置に通知することで、端末装置がデータ処理装置に対して負荷情報を問い合わせなくとも、各データ処理装置の負荷状況の変動が、端末装置自身が記憶している負荷情報に反映される。これにより、従来の方式に比べ、端末装置からの負荷情報の問い合わせは、端末装置の稼働開始時以外は発生しない分だけ、負荷情報交換のオーバーヘッドが削減される。また、負荷情報を管理しているデータ処理装置における負荷情報の問い合わせに係わる負荷も削減されるものである。

【図面の簡単な説明】

【図1】本発明の一実施例に係わる分散処理システムの構成を示すブロック図。

【図2】図1中の管理用データ処理装置20-1、データ処理装置20-2の詳細な構成を示すブロック図。

【図3】本実施例におけるデータ処理装置20-1~20-nの稼働開始時の動作を説明するためのフローチャート。

【図4】本実施例における管理用データ処理装置20-1がデータ処理装置20-1~20-nから負荷情報の通知を受けた場合の動作を説明するためのフローチャート。

【図5】本実施例における端末装置10-1~10-nの稼働開始時の動作を説明するためのフローチャート。

【図6】本実施例における管理用データ処理装置20-

1が端末装置10-1~10-nもしくはデータ処理装置20-2~20-nから負荷情報の問い合わせを受けた場合の動作について説明するフローチャート。

【図7】本実施例における端末装置10-1~10-nのデータ処理装置20-1~20-nに対して処理要求を発行する場合の動作について説明するためのフローチャート。

【図8】本実施例におけるデータ処理装置20-1~20-nが端末装置20-1~20-nから処理要求を受けた場合の動作を説明するための部分的なフローチャート。

【図9】本実施例におけるデータ処理装置20-1~20-nが端末装置20-1~20-nから処理要求を受けた場合の動作を説明するための部分的なフローチャート。

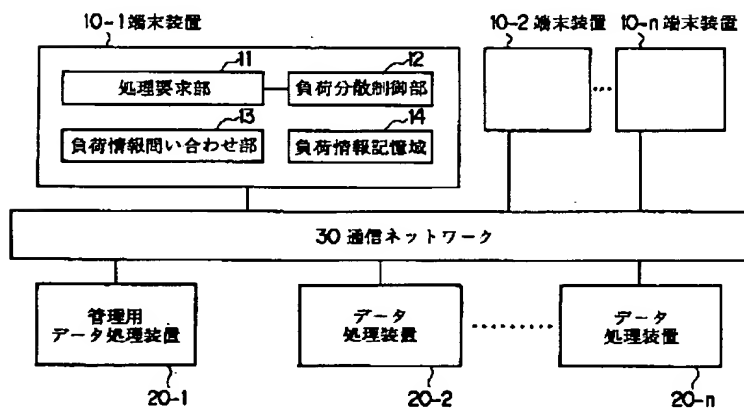
【図10】本実施例におけるデータ処理装置20-1~20-nが稼働時に随時実施する負荷状況監視の動作を説明するためのフローチャート。

【図11】従来の分散処理システムにおける負荷分散制御方式を説明するためのブロック図。

【符号の説明】

10-1~10-n…端末装置、11…処理要求部、12…負荷分散制御部、13…負荷情報問い合わせ部、14…負荷情報記憶域、20-1…管理用データ処理装置、20-2~20-n…データ処理装置、21…要求処理実施部、22…負荷状況検査部、23…負荷情報返送部、24…負荷情報問い合わせ部、25…負荷状況監視部、26…負荷情報通知部、27…負荷情報問い合わせ処理部、28…負荷情報通知処理部、29…負荷情報管理域、30…通信ネットワーク。

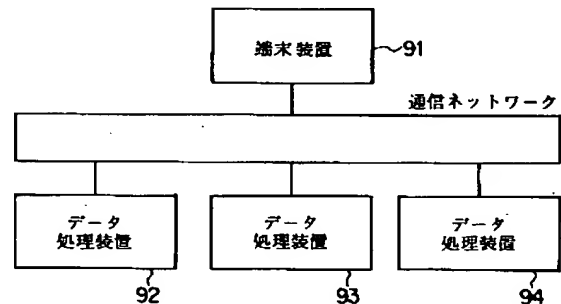
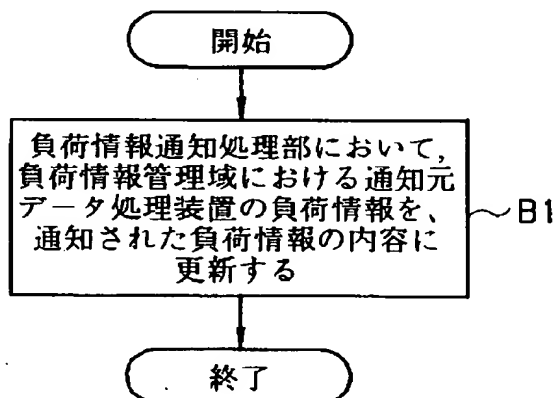
【図1】



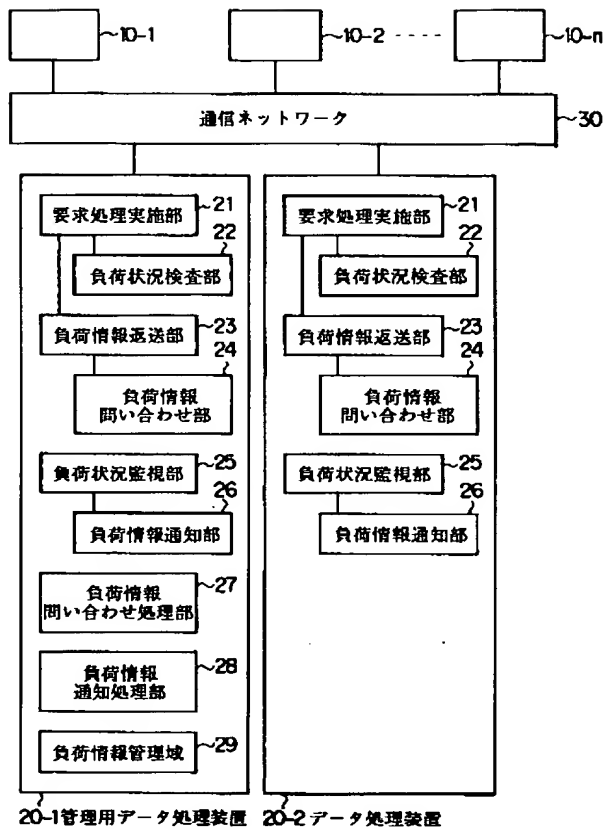
【図4】

【図11】

管理用データ処理装置において、データ処理装置から、負荷情報の通知を受けた場合の処理

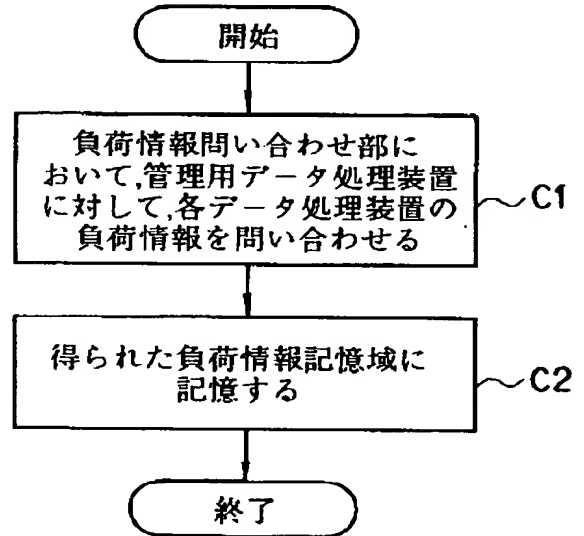


【図 2】



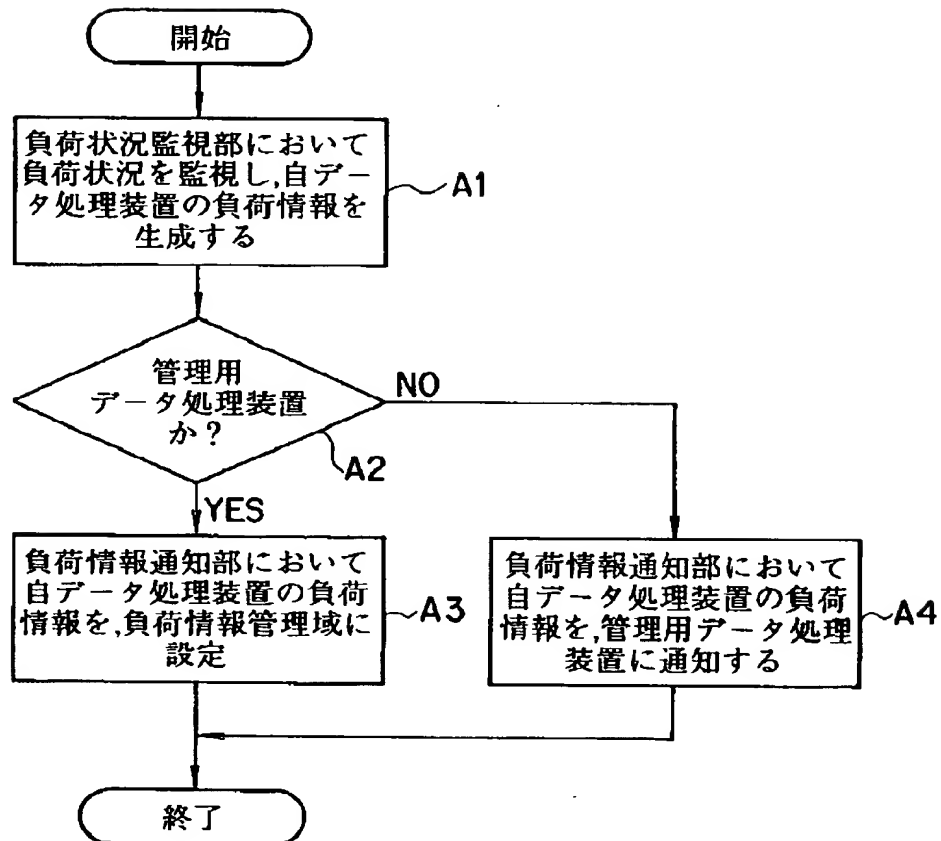
【図 5】

端末装置における、稼働開始時の処理



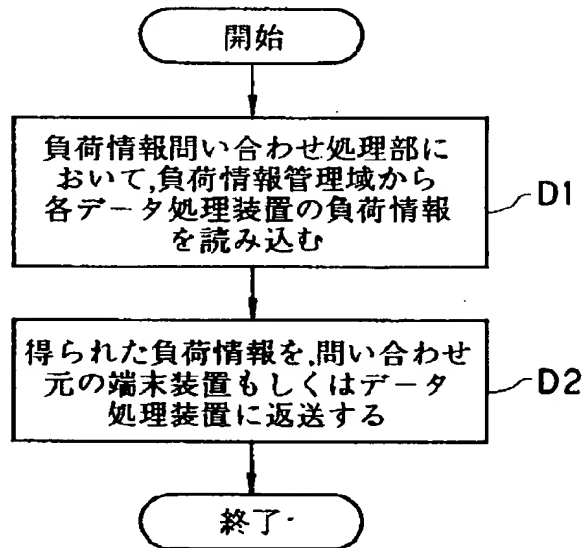
【図3】

データ処理装置における稼働開始時の処理



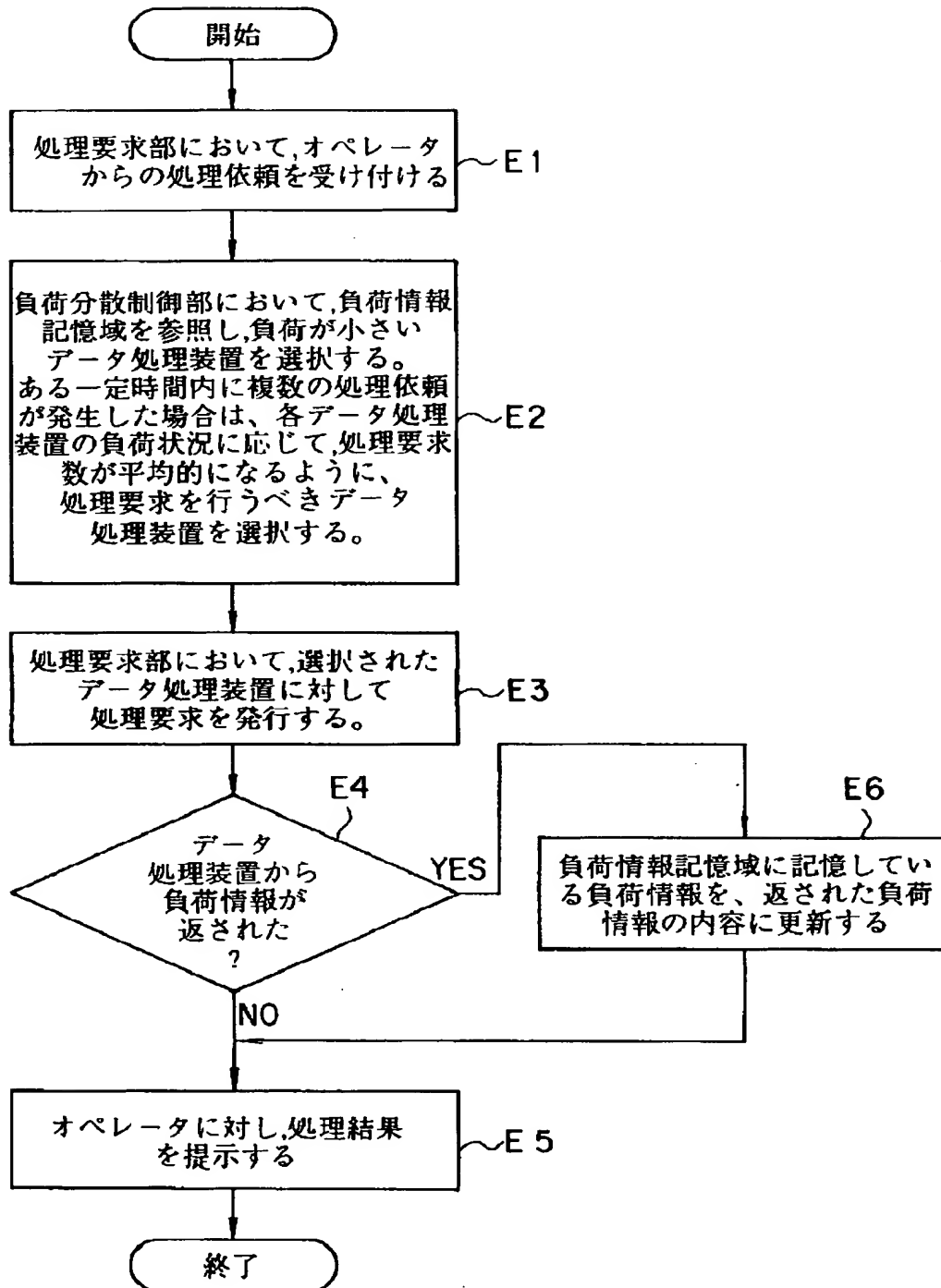
【図6】

管理用データ処理装置において、端末装置もしくはデータ処理装置から、負荷情報の問い合わせを受けた場合の処理



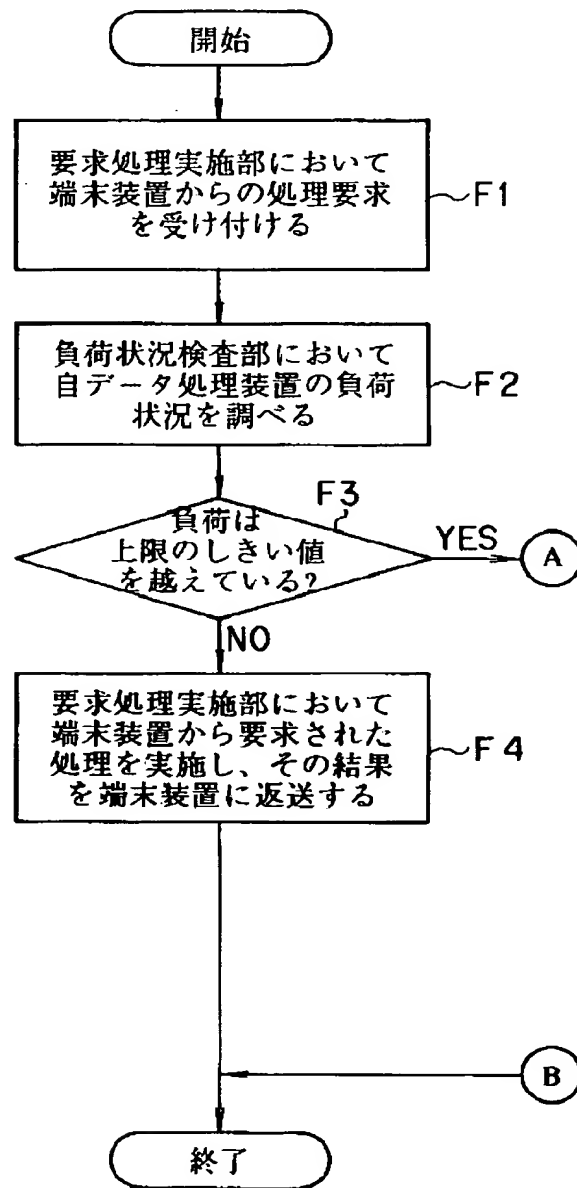
【図 7】

端末装置における、処理要求時の処理



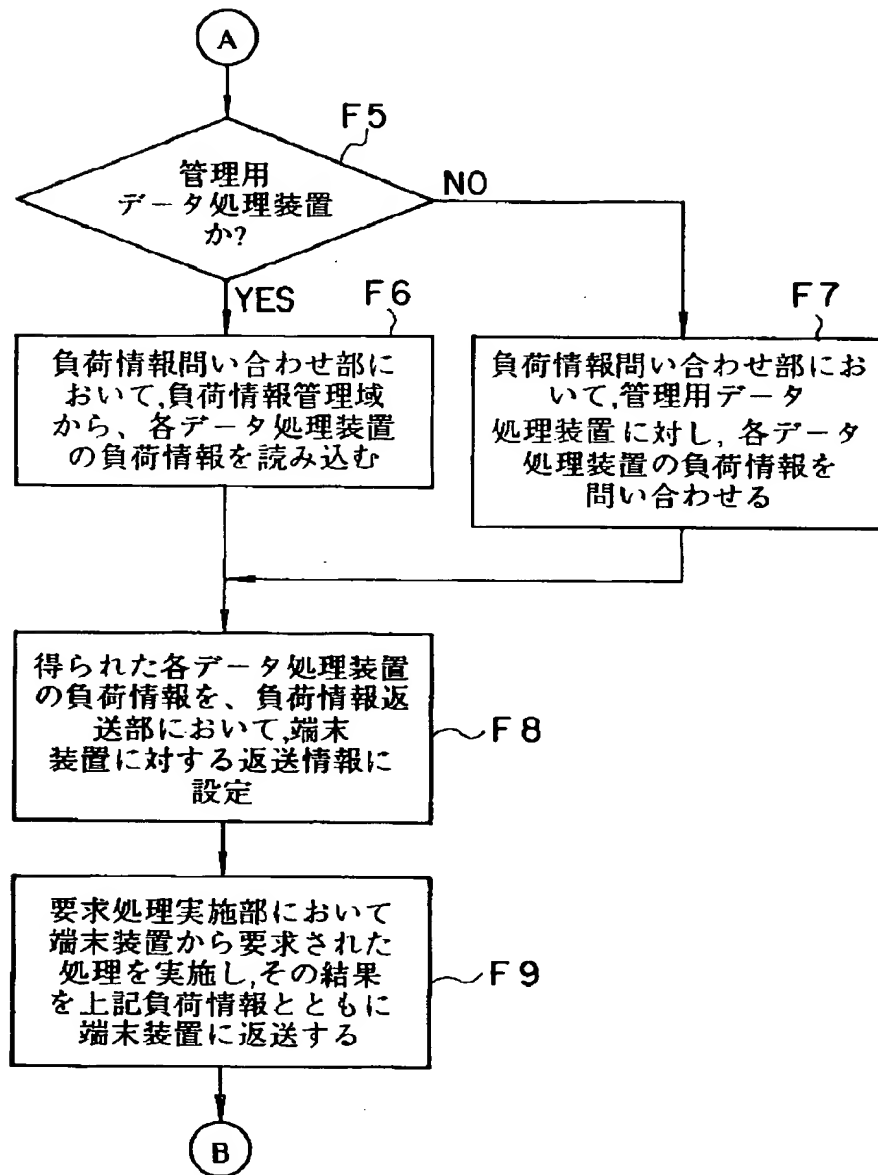
【図 8】

データ処理装置における、端末装置からの処理要求受付時の処理(その1)



【図 9】

データ処理装置における、端末装置からの処理要求受付時の処理(その2)



【図10】

データ処理装置における、稼働時の負荷状況監視処理

